

FIGURE 1

The Universal Code Equivalent of
the Mitochondrial 1-Sce I Gene.

AAAAAATAAAATCAT ATG AAA AAT ATT AAA AAA AAT CAA GTA ATG AAT CTC GGT GGT AAT TCT
 M K N I K K N Q Y M H L Q P H S
 AAA TTA TTA AAA GAA TAT AAA TCA CAA TTA ATT GAA TTA AAT ATT GAA CAA TTT GAA GAA
 K L L K E Y K S Q L I E L M I E Q F E A
 GGT ATT GGT TTA ATT TTA GGA GAT GGT TAT ATT CGT AGT CGT GAT GAA GGT AAA ACT TAT
 G I G L I L G D A Y I R S R D E S K T Y
 TGT ATG CAA TTT CAC TCC AAA AAT AAG GCA TAC ATG GAT CAT GTA TGT TTA TTA TAT GAT
 C M Q F E W K N K A Y M D H V C L L Y D
 CAA TGG GTA TTA TCA CCT CCT CAT AAA AAA GAA AGA GTT AAT CAT TTA GGT AAT TTA GTA
 Q W Y L S P P H K K E R V M H L G N L V
 ATT ACC TGG GGA GGT CAA ACT TTT AAA CAT CAA GGT TTT AAT AAA TTA GGT AAC TTA TTT
 I T W C A Q T F K M Q A F M K L A H L F
 ATT GTA AAT AAT AAA AAA CTT ATT CCT AAT AAT TTA GTT GAA AAT TAT TTA ACA GGT ATG
 I Y N M K K L I P H N L V E N Y L T P H
 AGT CTG GCA TAT TGG TTT ATG GAT GAT GGA GGT AAA TGG GAT TAT AAT AAA AAT TCT GGT
 S L A Y W F M D D Q G K W D Y H K H S L
 AAT AAA AGT ATT GTA TTA AAT ACA CAA AGT TTT ACT TTT CAA GAA GTA GAA TAT TTA GGT
 M K S I V L N T Q S F T F E E V C Y L V
 AAA GGT TTA AGA AAT AAA TTT CAA TTA AAT TGT TAT GTT AAA ATT AAT AAA AAT AAA CCA
 K G L R N K F Q L M C Y Y K I M K N K P
 ATT ATT TAT ATT GAT TCT ATG AGT TAT CTG ATT TTT TAT AAT TTA ATT AAA CCT TAT TTA
 I I Y I D S H S Y L I F Y H I I K P Y L
 ATT CCT CAA ATG ATG TAT AAA CTG CCT AAT ACT ATT TCA TCC GAA ACT TTT TTA AAA TAA
 I P Q M M Y K L P H T I S S E T F L K

6640220" 06444260

FIGURE 2

The synthetic I-Sce I gene

Bam HI
|

CCGGATCCATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GGT CCG AAC TGT
M H M K N I K K N Q Y M N L S P N S

AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA
K L L K E Y K S Q L I E L N I E S F E A

1. GGT ATC GGT CTG ATC CTG GGT GAT GGT TAC ATC CGT TCT CGT GAT GAA GGT AAA ACC TAC
G I G L I L G D A Y I R S R N E S K T Y

TGT ATG CAG TTC CAG TGG AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT
C M Q F E W K N K A Y M D H V C L L Y D

CAG TGG GTA CTG TCC CCG CCG CAC AAA AAA CAA CGT GTT AAC CAC CTG GGT AAC CTG GTA
Q W V L S P P H K K E R Y N H L S N L Y

ATC ACC TGG GGC GGC CAG ACT TTC AAA CAC CAA GCT TTC AAC AAA CTG GCT AAC CTG TTC
I T W G A Q T F K M Q A F N K L A N L F

ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA AAC TAC CTG ACC CCG ATG
I Y N N K K T I P N N L Y E N Y L T P M

2. TCT CTG GCA TAC TGG TTC ATG GAT GAT GGT GGT AAA TGG GAT TAC AAC AAA AAC TCT ACC
S L A Y W F M D D G G X W D Y N K N S T

AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GGT
N K S I V L N T Q S F T F E E V E Y L Y

AAG GGT CTG CGT AAC AAA TTC CAA CTG AAC TGT TAC GTA AAA ATC AAC AAA AAC AAA CCG
X G L R N K F Q L N C Y Y K I H K N K P

ATC ATC TAC ATC GAT TCT ATG TCT TAC CTG ATC TTC TAC AAC CTG ATC AAA CCG TAC CTG
I I Y I D S M S Y L I F Y N L I K P Y L

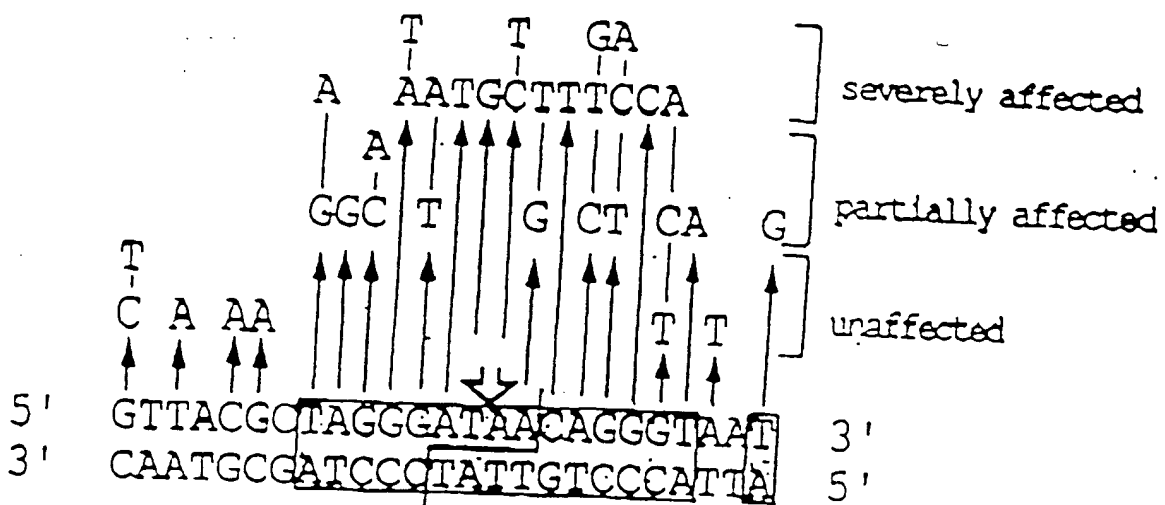
ATC CCG CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCG TCG GAA ACT TTC CTG AAA TAA
I P Q M M Y K L P N T I S S E T F L K

TAAGTCGACTGCAGGATCCGGTAAGTAAGTAA
| | |
Sall PstI BamHI

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.

FIGURE 3

66000" 06711200



Introduction

Selfish

I-SceI coding sequence of pSCM525 - Note the two amino acid N-terminal extension as compared to genuine version of the gene.

VARIATIONS AROUND THE 1-Scp 1 SEQUENCE

	-2	-1	1	2	3	4	5	6	7	8	9	10							
	M	H	V	Y	V	I	V	X	V	I	V	N	V	L	S	P	V	S	
			20									30							
K	L	L	K	E	Y	K	S	D	L	I	E	L	N	I	E	S	F	E	A
			40									50							
G	I	S	L	I	L	A	A	Y	I	R	S	R	D	E	S	K			
			60									70							
C	H	D	F	E	H	K	H	K	A	Y	H	D	H	V	S	L			
			80									90							
Q	V	Y	L	S	P	P	H	K	K	E	R	Y	N	H	L	A	H		
			100									110							
I	T	V	G	A	D	T	F	K	H	D	A	F	N	K	L	A	N		
			120									130							
I	V	N	H	K	K	I	I	P	N	N	L	V	E	H	Y	L	T	D	
			140									150							
D	L	A	Y	V	F	H	D	D	S	S	K	V	D	Y	H	K	H	S	
			160									170							
H	K	S	I	V	L	N	T	D	S	F	T	F	E	E	V	E	Y		
			180									190							
K	G	L	R	N	K	F	D	L	N	C	Y	V	K	I	H	K	N	K	D
			200									210							
I	I	Y	I	D	S	H	S	Y	L	I	F	Y	H	L	I	K	P	Y	
			220									230							
I	P	Q	H	H	Y	K	L	P	H	T	I	S	S	E	T	F	L	K	

Positions that can be changed without affecting enzyme activity (demonstrated)

positions -1 and -3 are not natural. The two amino acids are added due to cloning strategies

positions 1 to 10: can be deleted

position 36: G is tolerated

position 40: M or Y are tolerated

position 41: S or H are tolerated

position 43: A is tolerated

position 46: Y or N are tolerated

position 91: A is tolerated

positions 153 and 158: L are tolerated

position 223: A and S are tolerated

Changes that affect enzyme activity (demonstrated)

position 19: L to S

position 38: I to S or N

position 39: G to D or R

position 40: L to Q

position 42: L to R

position 44: D to E, G or H

position 45: A to E or D

position 46: Y to D

position 47: I to R or Y

position 50: L to S

position 144: D to E

position 145: D to E

660300"06TH200

Group I Intron Encoded Endonucleases and Related Endonucleases

Endonuclease		Recognition Sequence	CLAY/CA Site	▼ Intron site
Two Dodecapptide Family (for 4 bp cutters)	I-Sce I (<i>Saccharomyces mitochondria</i>)	CGCTAGGGGATAA CAGGGGTAATATAGC GCGATCCCTATATATATATATATATATAT		
	I-Sce IV (<i>Saccharomyces mitochondria</i>)	TTCTCATGATCA AGCTTAAATCCATAG AAGAGTAC TAATCAGATAGGATAGC		
	I-Sce II (<i>Saccharomyces mitochondria</i>)	CTTTGGTCACTCC AGAAAGTAATATATAT GAATCCGAGTATATATATATATATATATAT		
	I-Ceu I (<i>Chlamydomonas chloroplast</i>)	TAACGGTCCCTAA GTTAGCCGAAATTTCA ATTGCCAG GATTCGATCGGCTTAAAT		
	I-Hpo I (<i>Physarum nucleus</i>)	TGACCTCTCTTAA GGTAGCCAAATCCCT ACTGAGCAGTATATATATATATATATATAT		
	I-Sce III (<i>Saccharomyces mitochondria</i>)	CGAGGTTTTTGTAACTATTTATTTAGC CCTCCGAAAGCATTGATTAATTAATAGC		
	I-Cre I (<i>Chlamydomonas chloroplast</i>)	GGGTTCAAAAACCTCGTGGAGACAGTTCT CCCAGTTTTTCAGTCATCTCTCTCTCTCT		
	Endo Sce I (RF3) (<i>Saccharomyces mitochondria</i>) (Not in series)	GATGCTGATAGC ATAGGCTTTATTTAT GTACGACATCTCTTATCTGAAACCAATA		
	HO (<i>Saccharomyces nucleus</i>) (Not in series)	CTTCCGCGC ACATATATAATTTTATTA TAAAGCTCTCTCTCTCTCTCTCTCTCTCT		
	I-Cem I (<i>Chlamydomonas mitochondria</i>) (Primitive endonuclease)	ACCATGGGGTCAAATGCTTTTCTGCG TGCTACCCCGAGTTTACAGAAAGACCC		
	I-Pan I (<i>Podosporea mitochondria</i>) (Putative endonuclease)	CGCCCTGAATGCA TATATATATATATAT CAGCCACTTACTATATAATATATATATAT		
Other Structural Families	(Bacteriophage T4)			
	I-Tev I	GAAAG CCCTCAGTAGATTTTTCTTGGCTCTACCGTTTAA GTTCCGAGTCACTTACAAAAGAACCCAGATGGCAATTA		
	I-Tev II	CAAGCTTATGATATGAAGTCAACACCGT TAT GTTGAAATACTCACTCTCACTTGTG CAATTA		
	I-Tev III	GCTATTCGTTCTT TATGTATCTTTTGGCTTATCTTTAA CCATTAAGCAAAATTCATAGAAAACCCACATGCAATAT		

FIGURE 7

EXPRESSION VECTORS

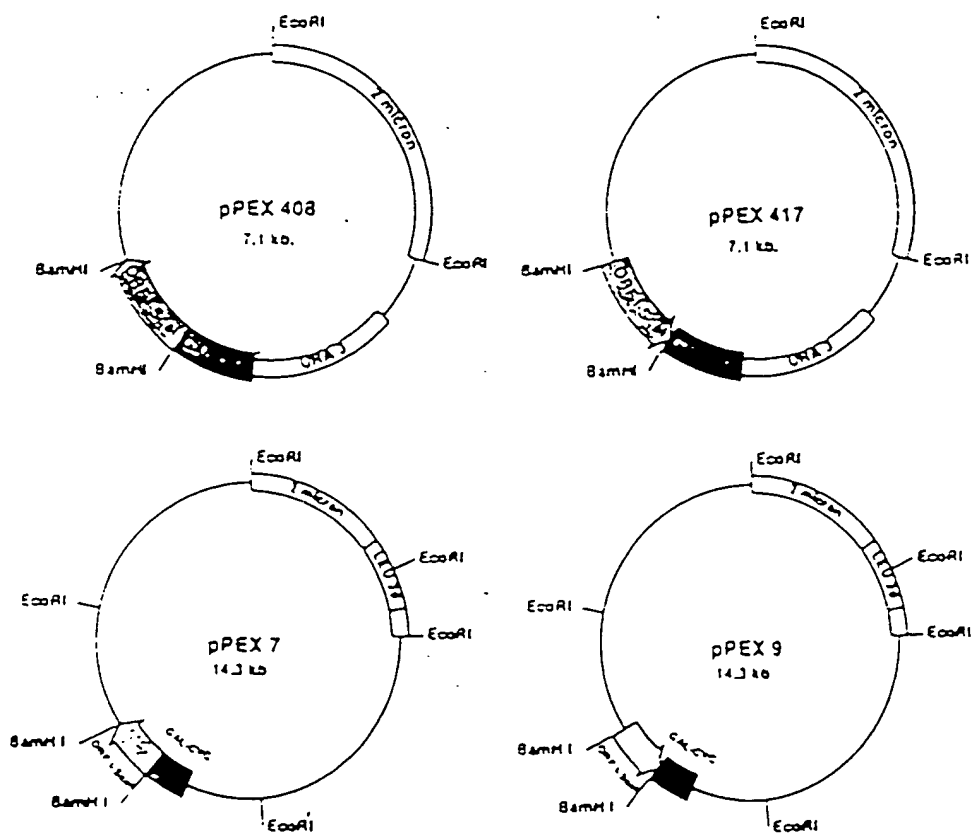
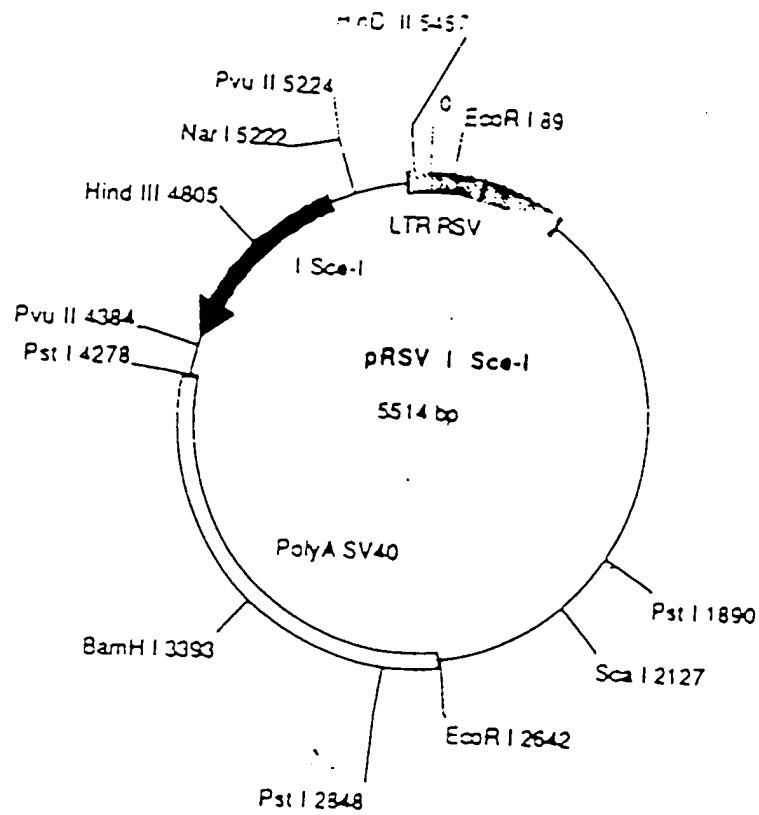
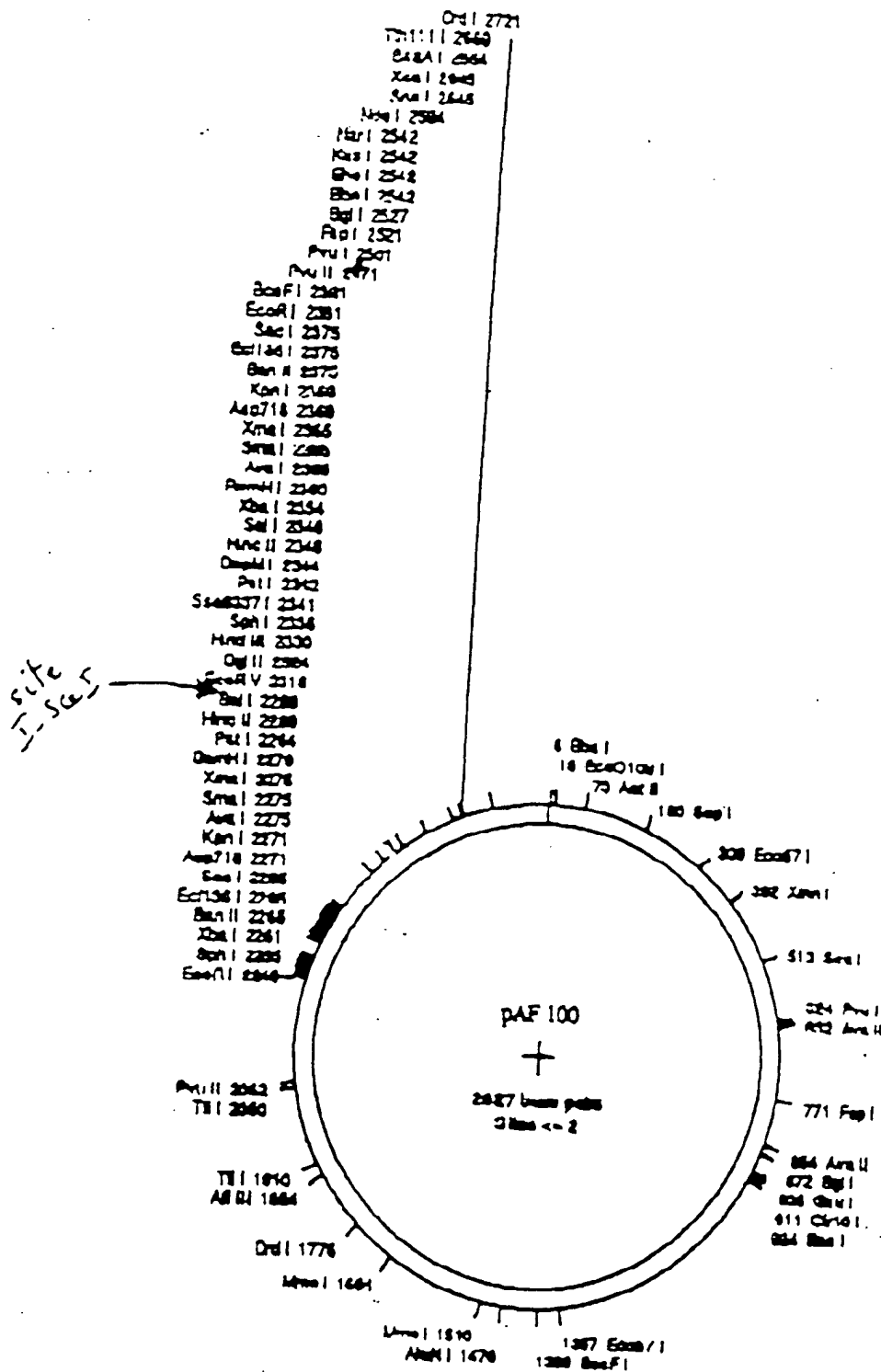


FIGURE 8



6013030"06711260

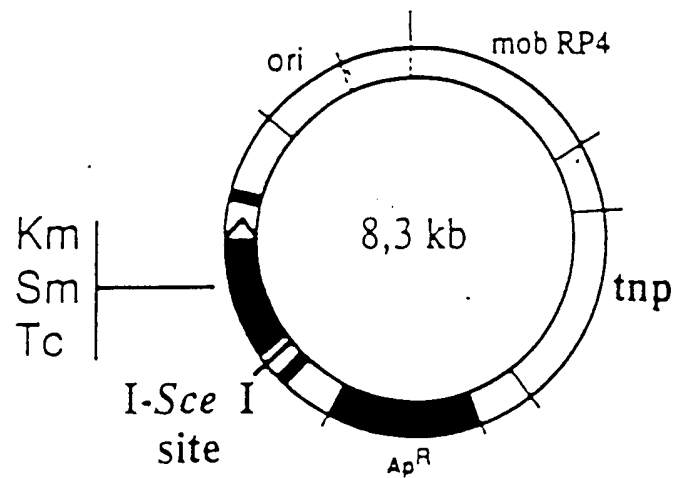
12345678910111213141516171819202122232425262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989910010110210310410510610710810911011111211311411511611711811912012112212312412512612712812913013113213313413513613713813914014114214314414514614714814915015115215315415515615715815916016116216316416516616716816917017117217317417517617717817918018118218318418518618718818919019119219319419519619719819920020120220320420520620720820921021121221321421521621721821922022122222322422522622722822923023123223323423523623723823924024124224324424524624724824925025125225325425525625725825926026126226326426526626726826927027127227327427527627727827928028128228328428528628728828929029129229329429529629729829930030130230330430530630730830931031131231331431531631731831932032132232332432532632732832933033133233333433533633733833934034134234334434534634734834935035135235335435535635735835936036136236336436536636736836937037137237337437537637737837938038138238338438538638738838939039139239339439539639739839940040140240340440540640740840941041141241341441541641741841942042142242342442542642742842943043143243343443543643743843944044144244344444544644744844945045145245345445545645745845946046146246346446546646746846947047147247347447547647747847948048148248348448548648748848949049149249349449549649749849950050150250350450550650750850951051151251351451551651751851952052152252352452552652752852953053153253353453553653753853954054154254354454554654754854955055155255355455555655755855956056156256356456556656756856957057157257357457557657757857958058158258358458558658758858959059159259359459559659759859960060160260360460560660760860961061161261361461561661761861962062162262362462562662762862963063163263363463563663763863964064164264364464564664764864965065165265365465565665765865966066166266366466566666766866967067167267367467567667767867968068168268368468568668768868969069169269369469569669769869970070170270370470570670770870971071171271371471571671771871972072172272372472572672772872973073173273373473573673773873974074174274374474574674774874975075175275375475575675775875976076176276376476576676776876977077177277377477577677777877978078178278378478578678778878979079179279379479579679779879980080180280380480580680780880981081181281381481581681781881982082182282382482582682782882983083183283383483583683783883984084184284384484584684784884985085185285385485585685785885986086186286386486586686786886987087187287387487587687787887988088188288388488588688788888989089189289389489589689789889990090190290390490590690790890991091191291391491591691791891992092192292392492592692792892993093193293393493593693793893994094194294394494594694794894995095195295395495595695795895996096196296396496596696796896997097197297397497597697797897998098198298398498598698798898999099199299399499599699799899910001001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110111021103110411051106110711081109111011111112111311141115111611171118111911201121112211231124112511261127112811291130113111321133113411351136113711381139114011411142114311441145114611471148114911501151115211531154115511561157115811591160116111621163116411651166116711681169117011711172117311741175117611771178117911801181118211831184118511861187118811891190119111921193119411951196119711981199120012011202120312041205120612071208120912101211121212131214121512161217121812191220122112221223122412251226122712281229123012311232123312341235123612371238123912401241124212431244124512461247124812491250125112521253125412551256125712581259126012611262126312641265126612671268126912701271127212731274127512761277127812791280128112821283128412851286128712881289129012911292129312941295129612971298129913



SECRET

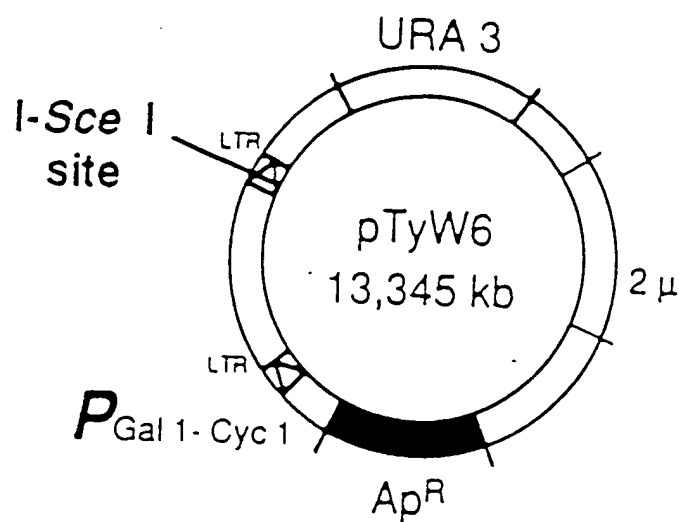
[illegible]

Names : pT ω Sm
 pT ω Km
 pT ω Tc



Construction : pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-SceI] in NotI unique site

FIGURE 12



Construction : pD 123 , from J. D. Boeke
with insertion of a linker [I-SceI - NotI] in BamHI

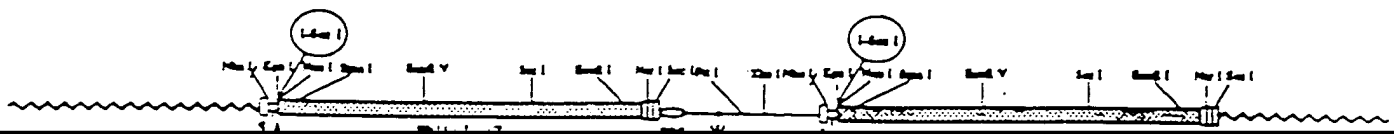
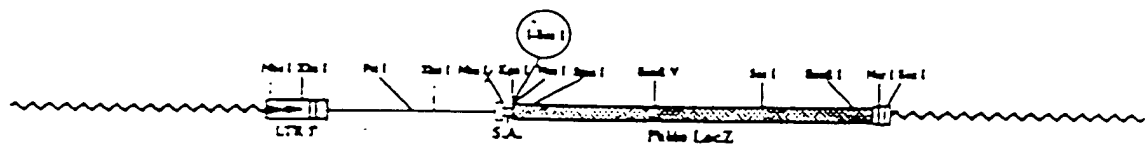
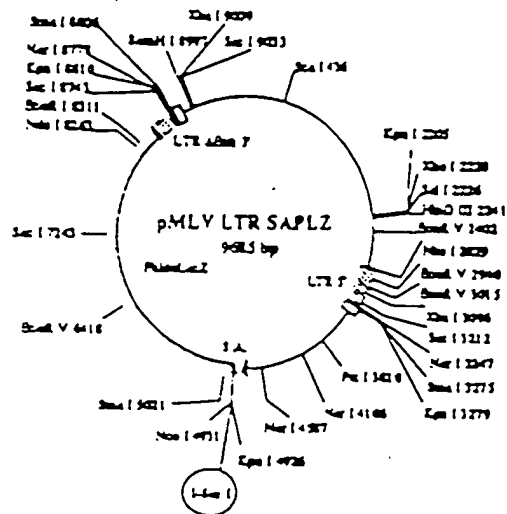


FIGURE 14

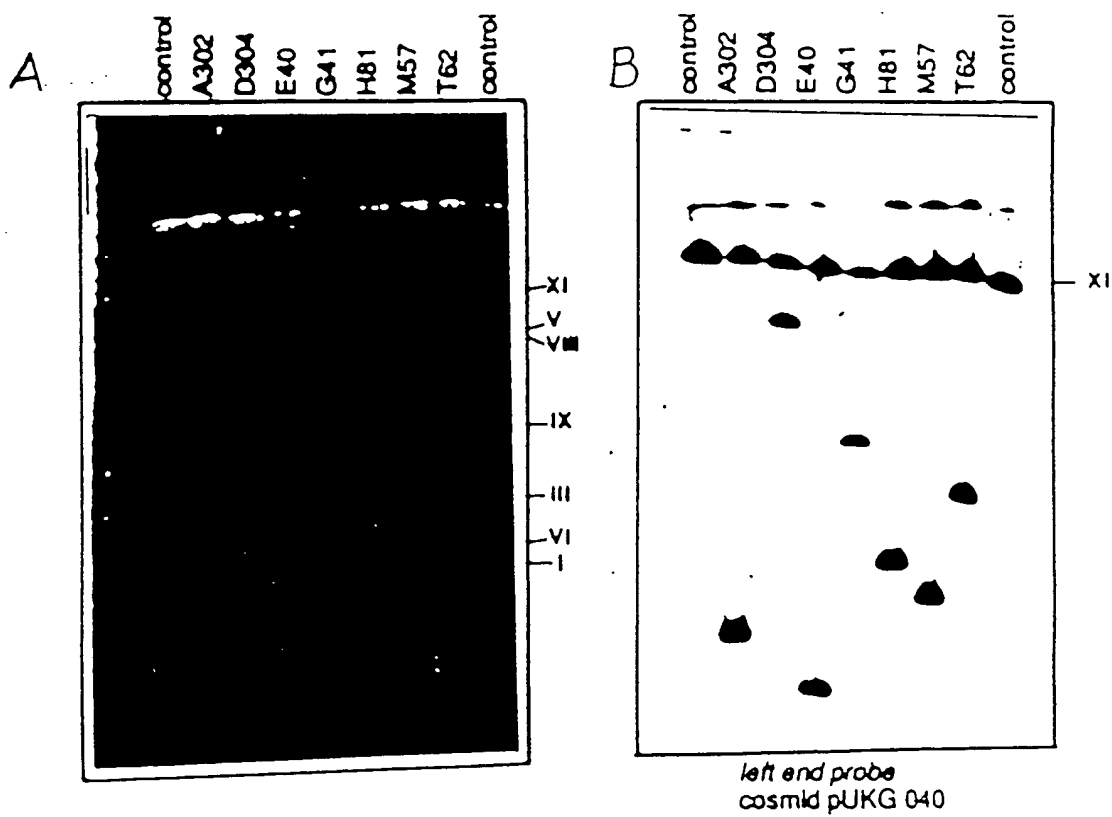


FIGURE 15

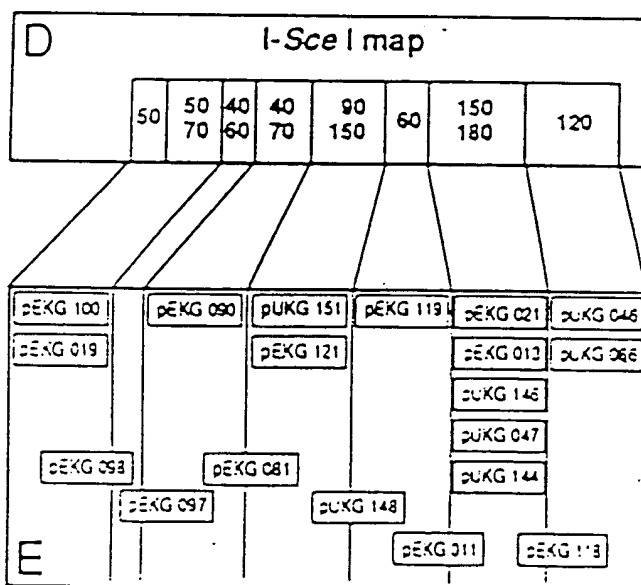
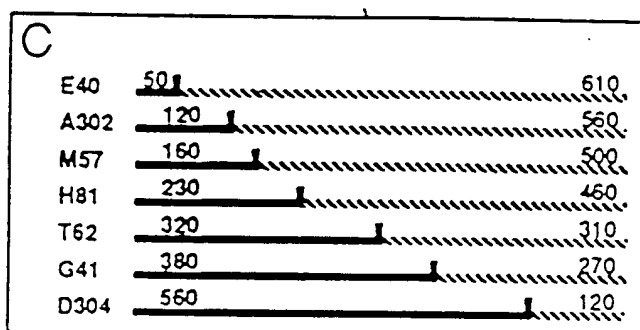
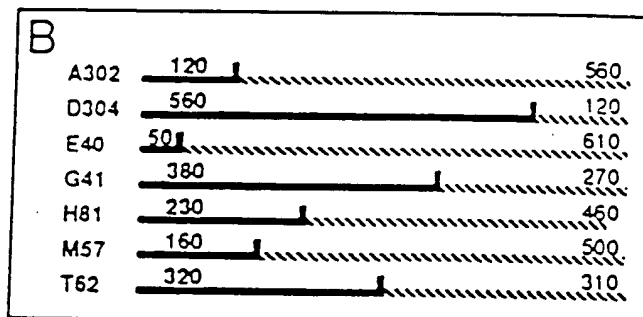
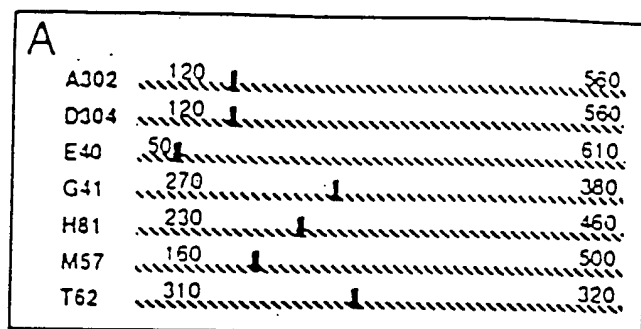


FIGURE 16

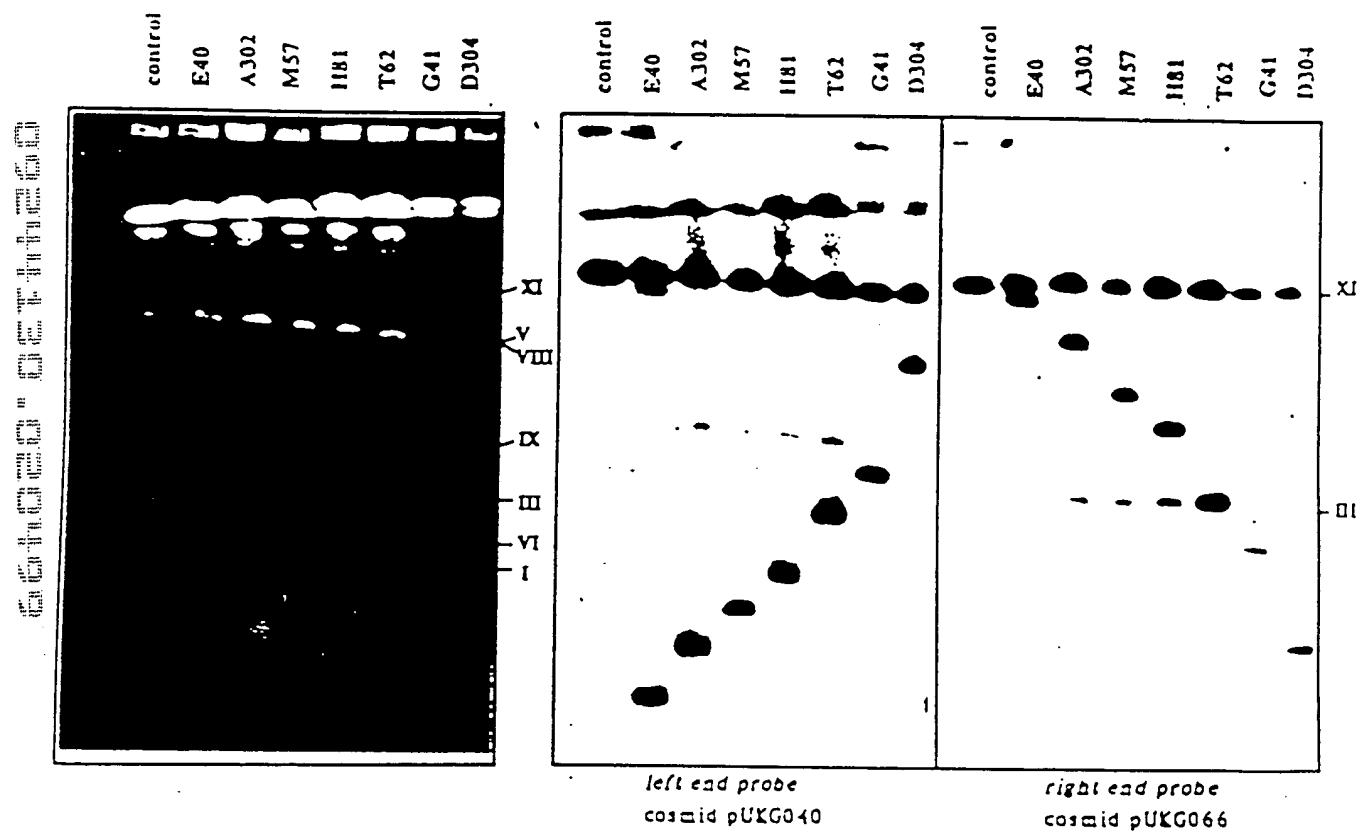
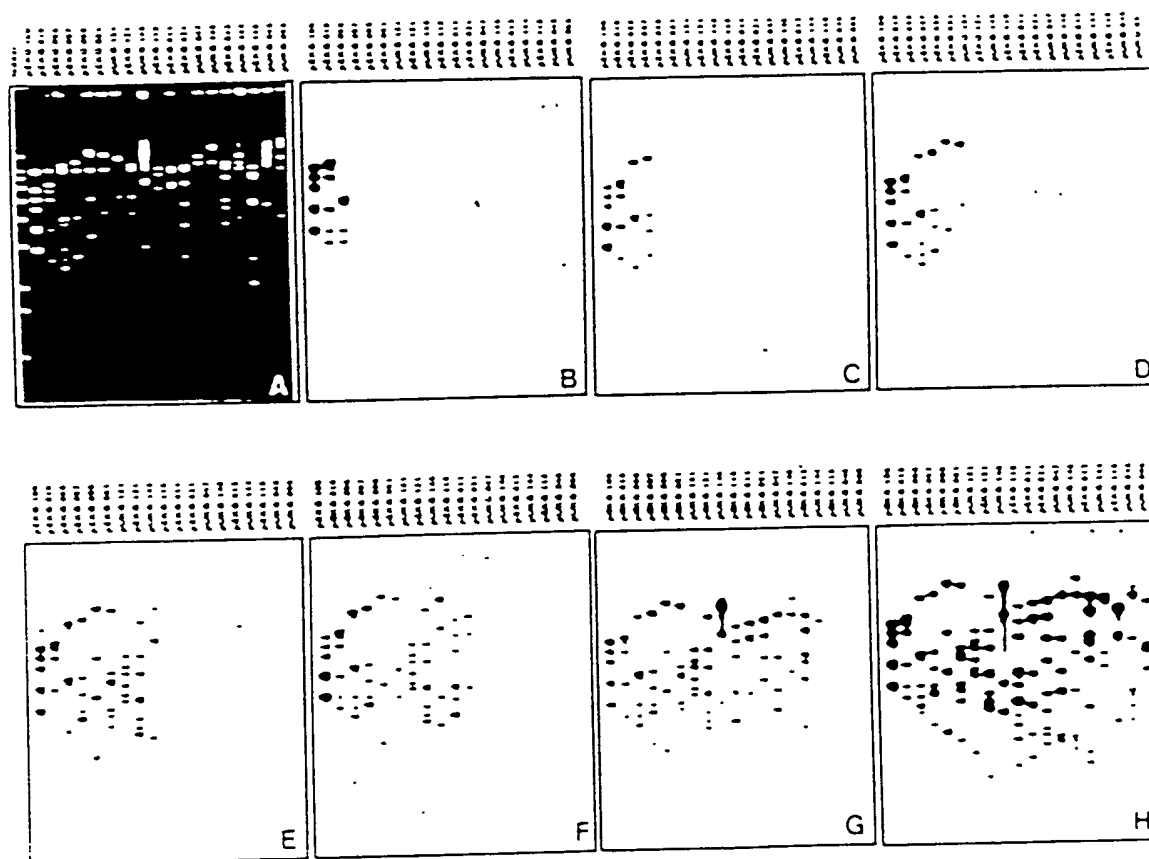
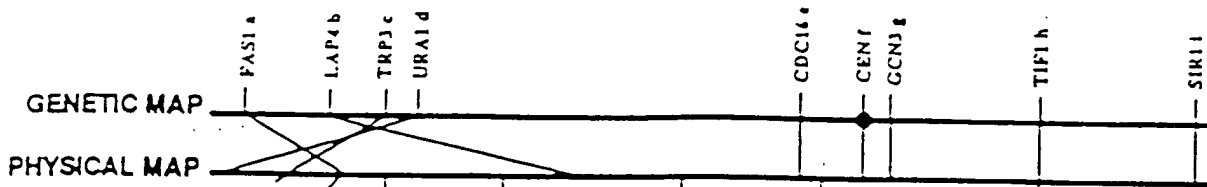


FIGURE 17





TEL	64 kb	64-78 kb	40-46 kb	44-78 kb	84-180 kb	80 kb	150-180 kb	120 kb	TEL
	pUKG 145 d	pEKG 025	pEKG 107	pEKG 006	pUKG 128 b	pEKG 002	pEKG 021	pUKG 064 h	
	pUKG 046 ad	pEKG 002	pEKG 007	pUKG 130	pEKG 002 b	pUKG 072	pUKG 128 i	pUKG 130	
	pEKG 012 ad	pUKG 129	pEKG 008	pEKG 024	pUKG 141 b	pUKG 144	pUKG 064 i	pEKG 104	
	pUKG 033 ad	pUKG 074	pEKG 007	pUKG 005	pUKG 151 b	pEKG 016	pUKG 047 i	pEKG 115	
	pUKG 077 ad	pEKG 010			pUKG 078 b	pUKG 001	pUKG 037 g	pEKG 116	
	pEKG 104 ad	pUKG 150			pEKG 006 b	pUKG 002	pEKG 008 g	pUKG 140	
		pUKG 136 a			pUKG 137 b	pEKG 004	pUKG 125 g	pEKG 120	
					pUKG 186 b	pUKG 064	pUKG 030 g	pEKG 064	
					pEKG 028	pEKG 008	pEKG 003 g	pEKG 100	
					pEKG 114	pEKG 110	pUKG 034 g	pUKG 152	
					pEKG 121	pUKG 003	pEKG 015 g	pUKG 046	
					pEKG 112	pUKG 022	pUKG 076 g	pUKG 052	
					pEKG 113	pEKG 110	pEKG 013 g	pUKG 131	
					pUKG 150	pEKG 027	pUKG 043	pUKG 136 i	
					pEKG 105	pEKG 014	pEKG 006	pEKG 101 i	
					pEKG 122	pEKG 008	pUKG 073	pEKG 007 i	
					pUKG 070	pUKG 007		pEKG 006 i	
					pEKG 020			pUKG 066 i	
								pUKG 123 i	
	pUKG 134 c								
	pEKG 023 c								
	pEKG 008 c								
	pEKG 001 c								
	pUKG 009 c	pEKG 065 a	pUKG 154	pEKG 001	pEKG 080	pEKG 011 e	pUKG 035 h		
	pEKG 017 c	pUKG 132 a	pUKG 057	pEKG 001	pUKG 135	pUKG 036 e	pUKG 124 h		
	pUKG 055 c	pEKG 008 a	pUKG 037	pEKG 108	pUKG 140	pEKG 111 e	pEKG 008 h		
	pUKG 051 c	pUKG 157 a	pUKG 053	pUKG 155	pUKG 133	pUKG 147 e	pEKG 106 h		
	pEKG 004 c	pUKG 075 a	pUKG 058	pUKG 040		pEKG 010 e	pEKG 103 h		
	pUKG 029 c	pEKG 003 a		pUKG 030		pUKG 041	pEKG 003 h		
	pUKG 143 c	pUKG 142 a		pUKG 048			pUKG 144 h		

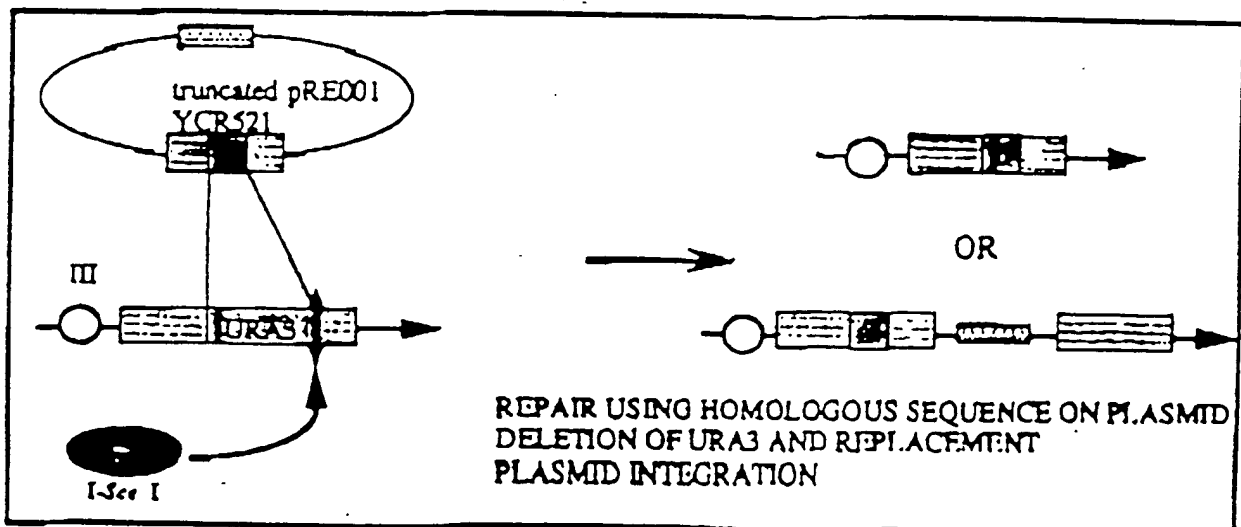
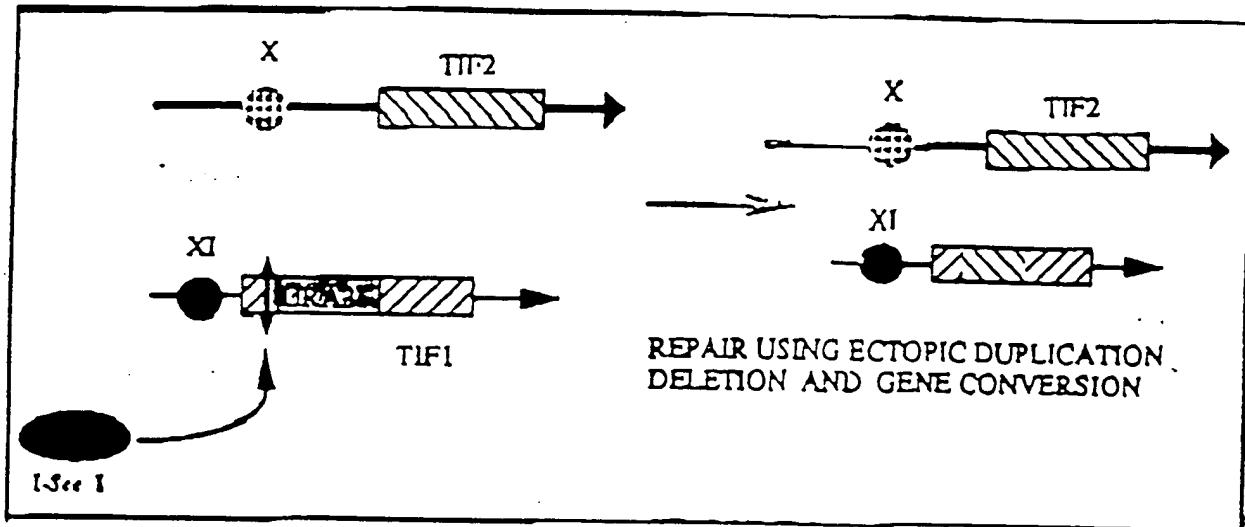


Figure 20

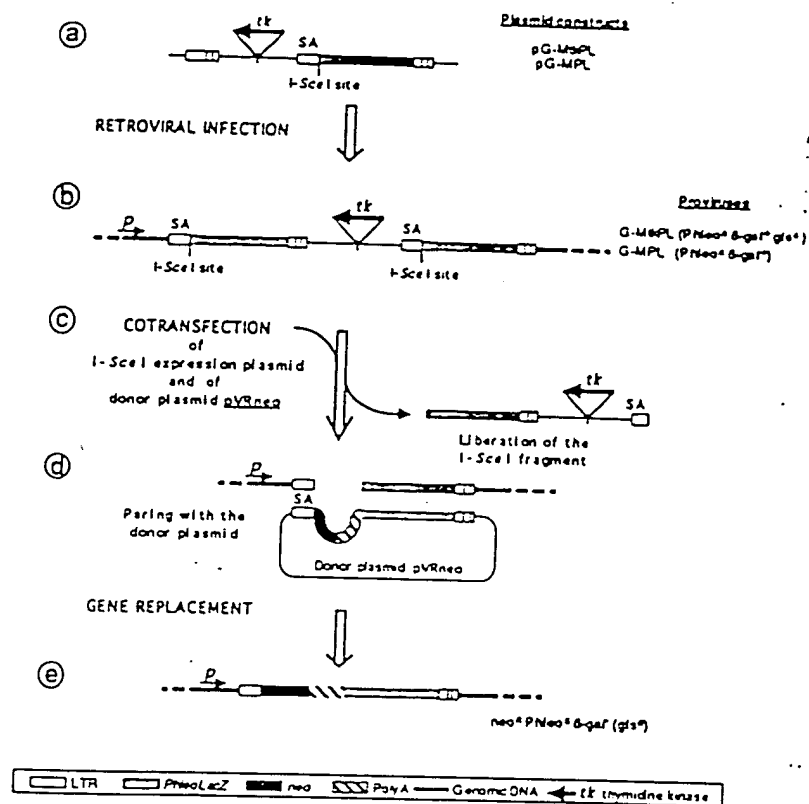


Figure 21

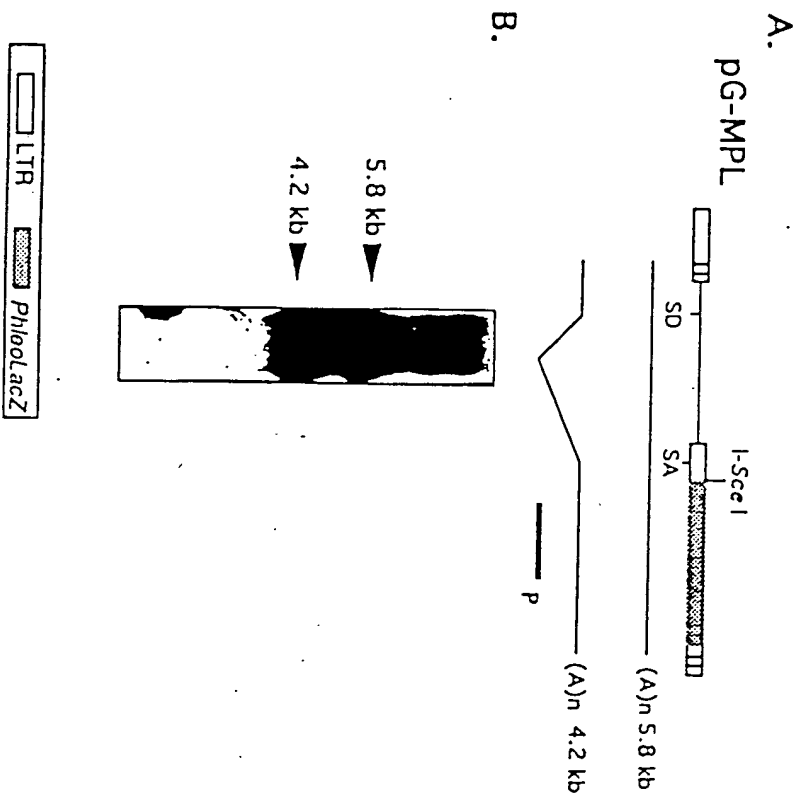
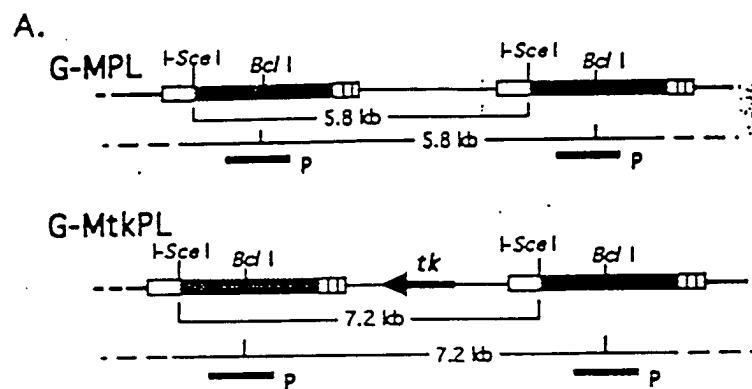
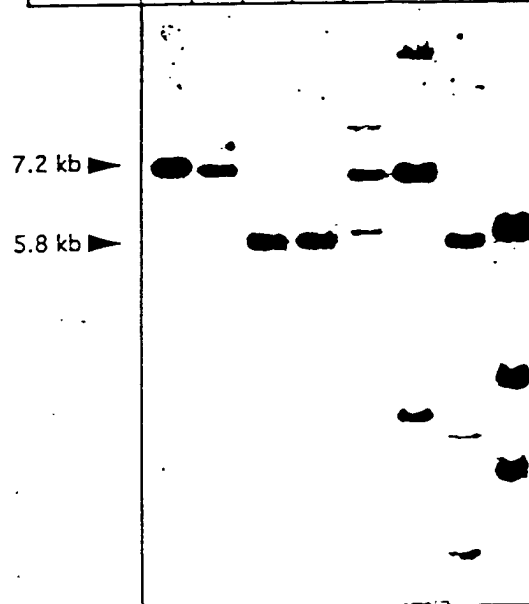
[illegible]

Figure 22



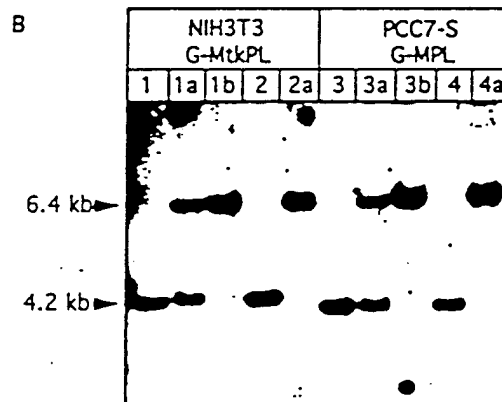
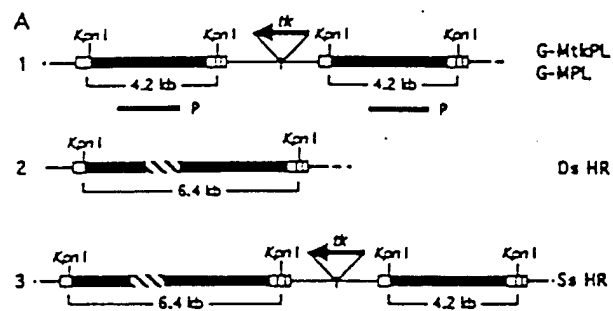
B.

ENZYME	I-Sce I				Bcl I			
CELL LINE	NIH3T3		PCC7-S		NIH3T3		PCC7-S	
PROVIRUS	G-MtkPL		G-MPL		G-MtkPL		G-MPL	
CLONE	1	2	3	4	1	2	3	4



LTR
 PhleoLacZ
 tk thymidine kinase

Figure 23



LTR
 PhleoLacZ
 neo
 PolyA
 Genomic DNA
 ← tk thymidine kinase

Figure 24

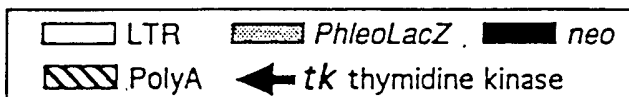
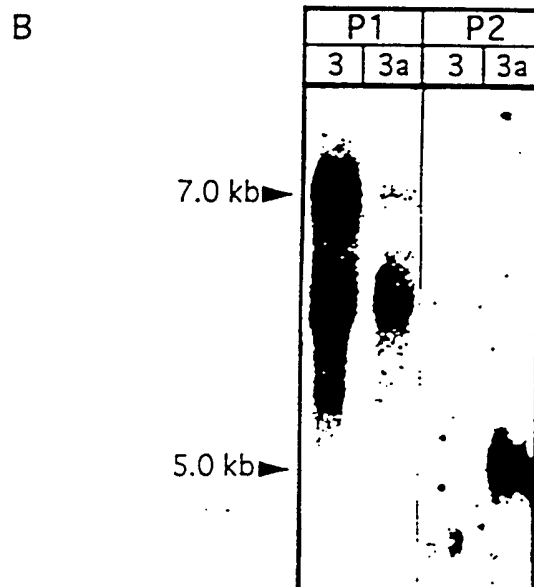
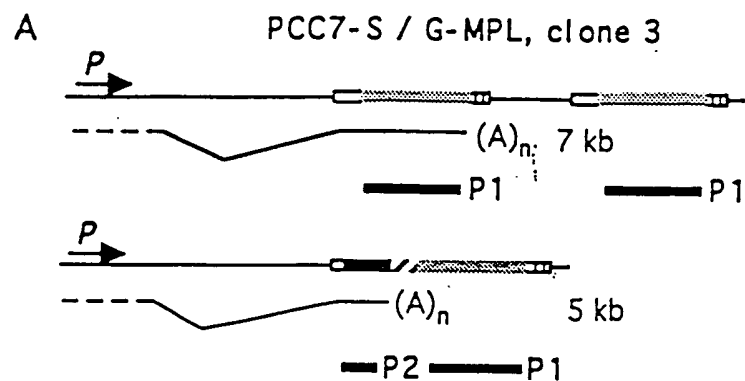
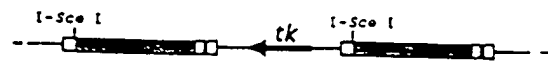


Figure 25

a. Chromosomal DNA containing provirus

Phenotypes

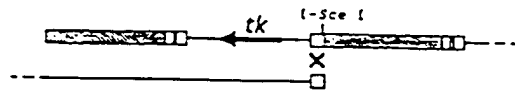


[Phleo^R, Glu^R, Δ-Gal⁺]

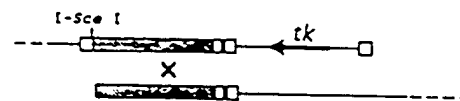
Transfection by
I-Sce I endonuclease
expression vector

b. Intra-chromosomal recombinations events

1. The left I-Sce I site is cut.
Pairing and recombination



2. The right I-Sce I site is cut.
Pairing and recombination



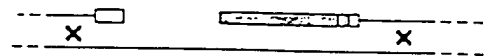
3. Both I-Sce I sites are cut.
Religation by end-joining



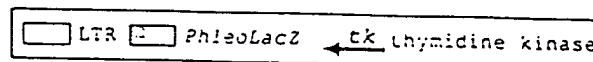
[Phleo^R, Glu^R, Δ-Gal⁺]

c. Inter-chromosomal recombination event

Both I-Sce I sites are cut. Gap repair using
intact chromosome sequences

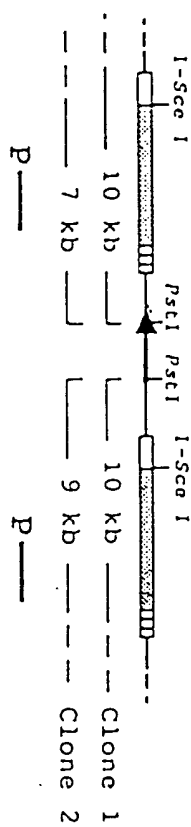


[Phleo^S, Glu^R, Δ-Gal⁻]



~~~~~~

a. Parental DNA, G-MtkPL



४.

| Part I |   |     |     |     |         |
|--------|---|-----|-----|-----|---------|
| ENZYME |   |     |     |     |         |
| CLONF. | 1 | 1.1 | 1.2 | 1.3 | 1.4 1.5 |



Q.

| ENZYME | Pat I |     |     |     |     |     |     |
|--------|-------|-----|-----|-----|-----|-----|-----|
| CLONE  | 2     | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 | 2.6 |



# 2025

— 22 —

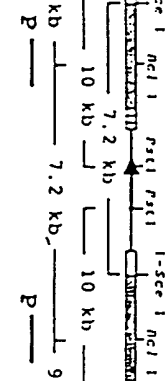
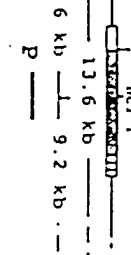


Diagram illustrating the structure of the pUC19 plasmid vector. The plasmid is circular, with a total size of 2600 bp. Key features include the origin of replication (ori) at 1000 bp, the lacZ' gene at 1300 bp, and a multiple cloning site (MCS) at 1300 bp. The MCS contains restriction sites for NotI, XbaI, SmaI, PstI, KpnI, SalI, BamHI, and EcoRI. The diagram also shows a 1.6 kb fragment and a 9.2 kb fragment, and a pUC19 label.



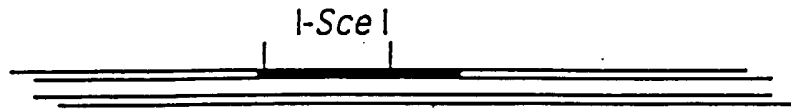
Q.

|        |   | pc1 1 |     |     |     |     |     |     |     |
|--------|---|-------|-----|-----|-----|-----|-----|-----|-----|
| CNTYPE |   |       |     |     |     |     |     |     |     |
| CLONE  | 2 | 2.a   | 2.b | 2.c | 2.d | 2.e | 2.f | 2.g | 2.h |

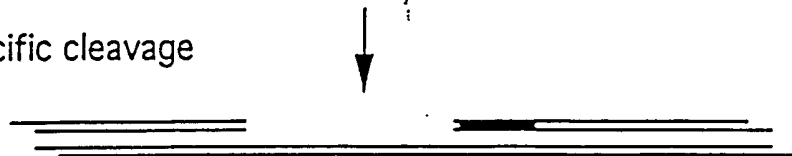


# LOSS OF HETEROZYGOSITY

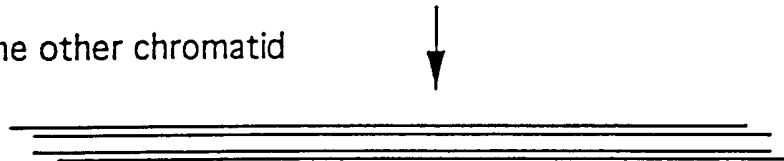
Integration of artificial site or presence of specific site



Expression of I-Sce I and specific cleavage



Repair of the DSB with the other chromatid



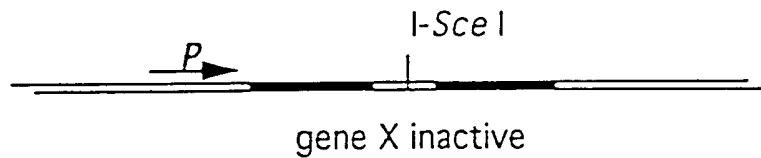
6610200" DET 11260

Figure 29

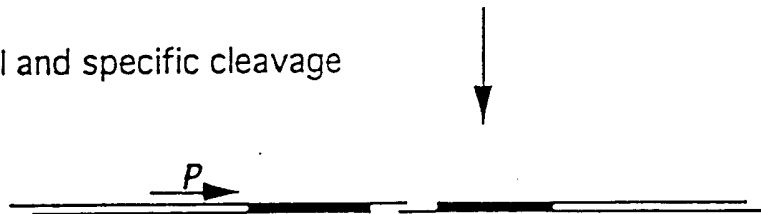
bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## CONDITIONAL ACTIVATION (Tandem repeat)

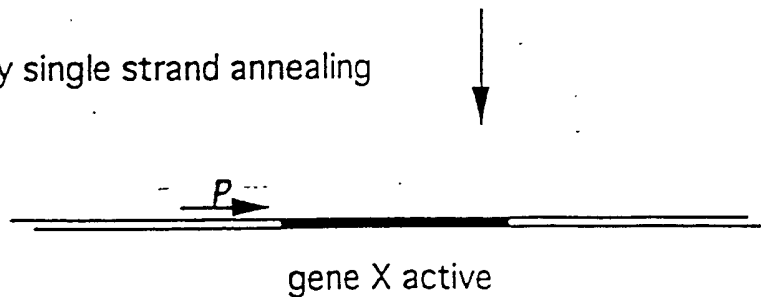
Integration of artificial site between tandem repeats



Expression of I-Sce I and specific cleavage



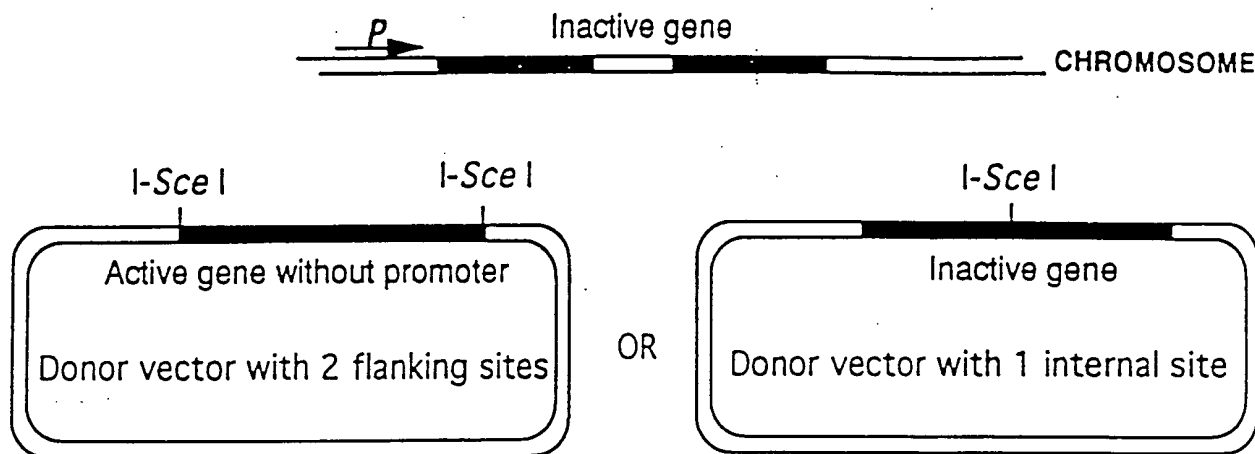
Repair of the DSB by single strand annealing



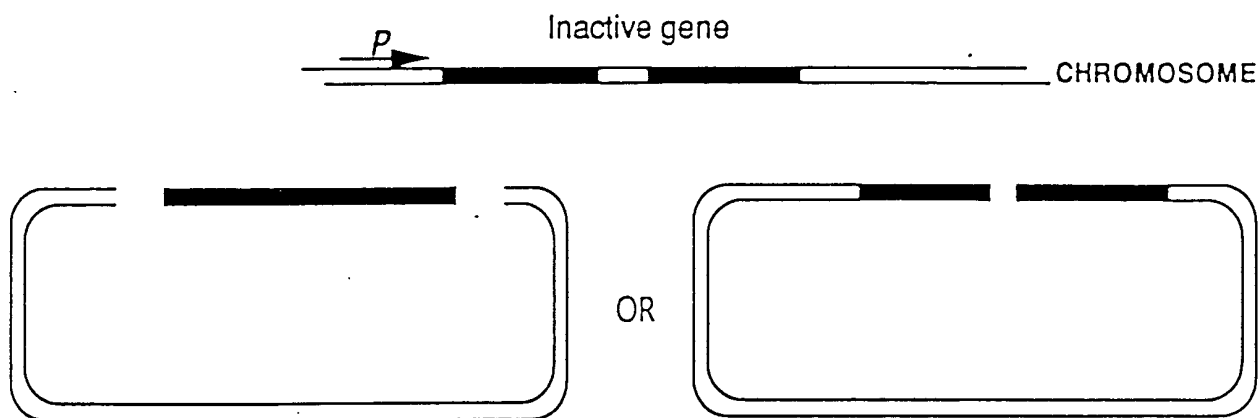
# ONE STEP REAGENTS

Figure 30

Integration of artificial site or presence of specific site



Expression of I-Sce I enzyme  
and  
specific cleavage of the donor plasmid



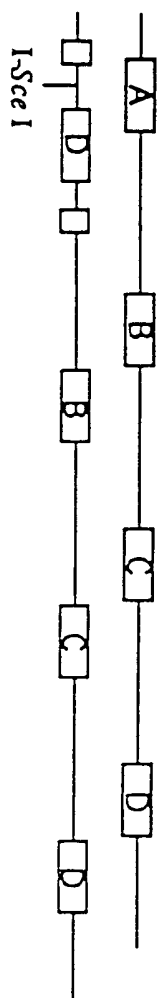
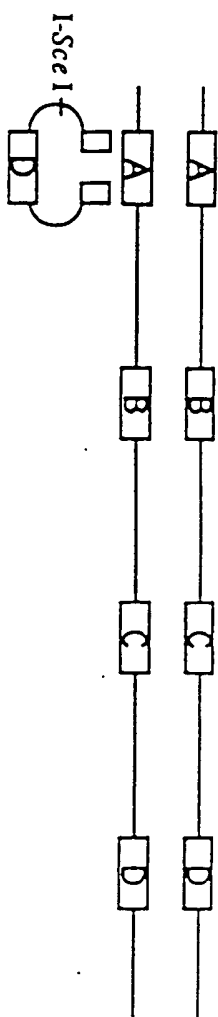
Recombination between the chromosome and plasmid



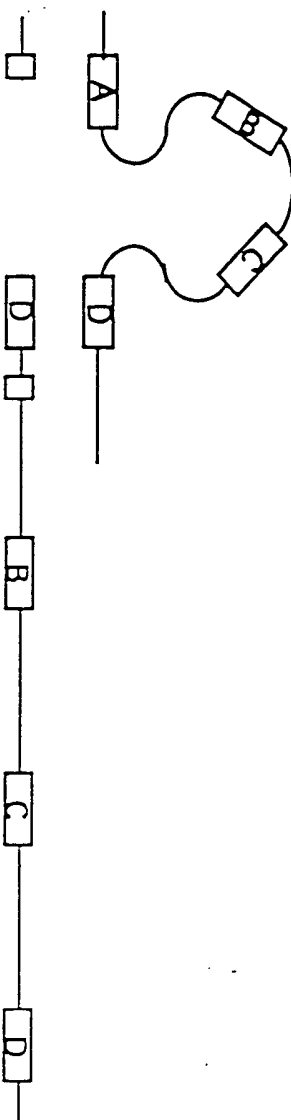
66000-001-1111

# DUPPLICATION OF A LOCUS

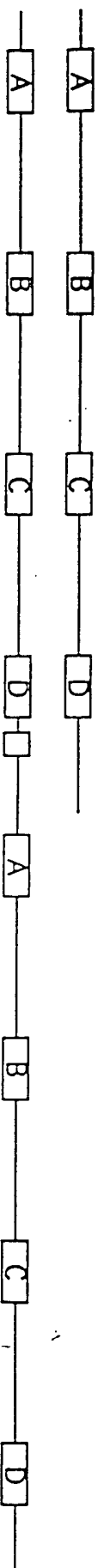
## 1 insertion of I-Sce I site by classical gene replacement



## 2 Specific cleavage by I-Sce I enzyme and repair of the break by homologous sequences



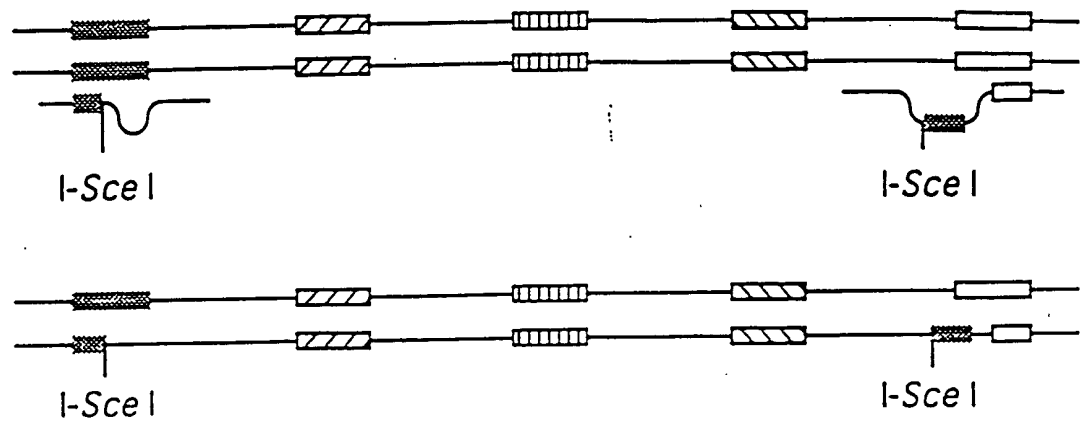
## 3 Duplication of the totality of the locus



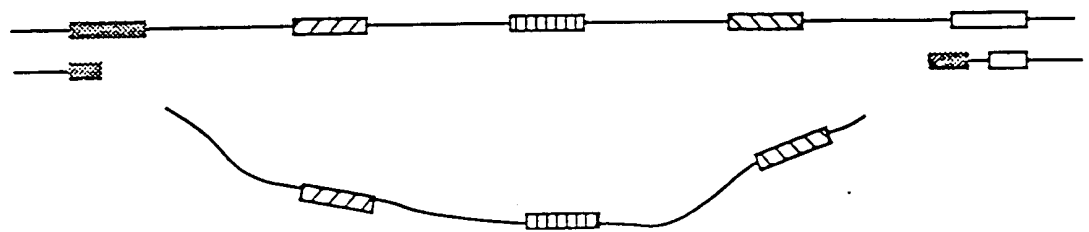


## DELETION OF A LOCUS

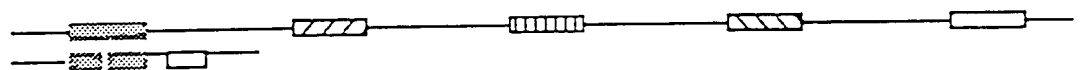
## 1 Insertion of two I-Sce I sites flanking the locus



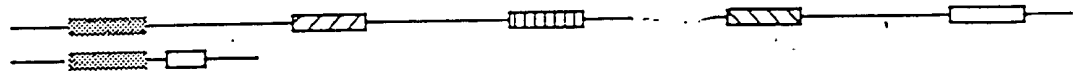
## 2 Expression of the enzyme and cleavage

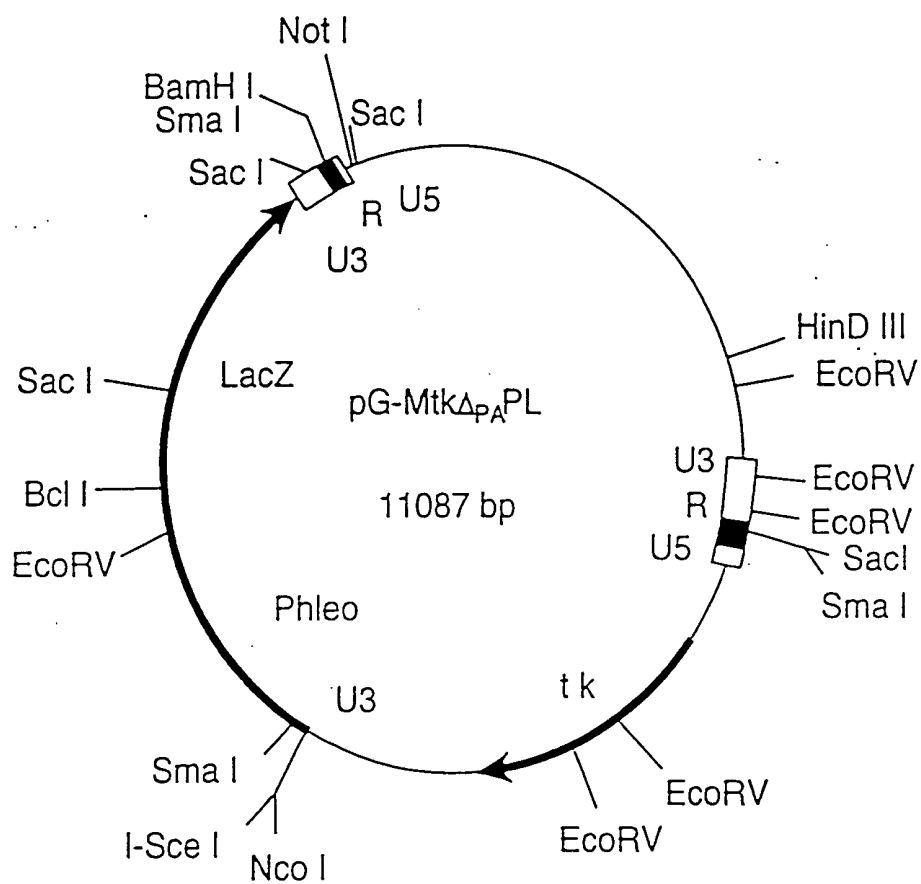


## 3 Recombination between the two ends



## 4 deletion of the locus





**FIG. 33**